INSECTICIDAL PROTEINS FROM PAECILOMYCES AND SYNERGISTIC COMBINATIONS THEREOF

The present invention relates to *inter alia*, insecticidal proteins and synergistic combinations thereof, DNA sequences encoding the proteins and methods of producing plants comprising said proteins and combinations. In particular the invention relates to insecticidal peptides obtainable from the fungi *Paecilomyces spp.*

Many fungi are pathogenic to insects, it is known that *Paecilomyces fumosoroseus* can be used as a biological control agent and this strain is sold commercially as a bio-control agent for use in greenhouses. *Hitherto* however, there has been no isolation and identification of insecticidal peptides from *Paecilomyces spp.*

It has been unexpectedly found that insecticidal peptides extracted from of *Paecilomyces spp.*, such as *Paecilomyces farinosus* provide a new type of potent orally active insecticidal peptide. Surprisingly, these proteins are also capable of acting synergistically with further proteins in particular CRY and VIP proteins.

According to the present invention there is provided an insecticidal protein comprising the sequence: X_1X_2ICTPAGVKCPAALPCCPGLRCIGGVNNKVCR (SEQ ID No. 1) wherein X_1 and X_2 are any amino acid. In a further embodiment of the present invention the amino acids at positions X_1 and X_2 are selected from the group consisting of: Glycine; Lysine; Serine; Tyrosine; Alanine; Methionine; Threonine; Glutamic acid; Aspartic acid; Asparagine and Valine. In a still further embodiment of the present invention the amino acids at positions X_1 and X_2 are Serine and Tyrosine respectively. In a still further embodiment of the present invention the amino acid a position X_1 is Glutamine. In a still further embodiment of the present invention the insecticidal protein comprises the sequence: GKICTPAGVKCPAALPCCPGLRCIGGVNNKVCR (SEQ ID No. 2). The present invention still further provides an insecticidal protein comprising the sequence: X_1X_2GKICTPAGVKCPAALPCCPGLRCIGGVNNKVCR (SEQ ID No. 3) wherein X_1 and X_2 are any amino acid, preferably an amino acid selected from the group consisting of: Glycine; Lysine; Serine; Tyrosine; Alanine; Methionine; Threonine; Glutamic acid; Aspartic acid; Asparagine and Valine, more preferably X_1 and X_2 are Serine and Tyrosine respectively.

The present invention further provides an insecticidal protein having at least 55% identity to any of the proteins depicted as SEQ ID Nos. 1 to 3. In a further embodiment of

the present invention the insecticidal protein has at least 60% identity to any of the proteins depicted as SEQ ID Nos. 1 to 3. In a still further embodiment of the present invention the insecticidal protein has at least 65% identity to any of the proteins depicted as SEQ ID Nos. 1 to 3. In a still further embodiment of the present invention the insecticidal protein has at least 70% identity to any of the proteins depicted as SEQ ID Nos. 1 to 3. In a still further embodiment of the present invention the insecticidal protein has at least 75% identity to any of the proteins depicted as SEQ ID Nos. 1 to 3. In a still further embodiment of the present invention the insecticidal protein has at least 80% identity to any of the proteins depicted as SEQ ID Nos. 1 to 3. In a still further embodiment of the present invention the insecticidal protein has at least 85% identity to any of the proteins depicted as SEQ ID Nos. 1 to 3. In a still further embodiment of the present invention the insecticidal protein has at least 90% identity to any of the proteins depicted as SEQ ID Nos. 1 to 3. In a still further embodiment of the present invention the insecticidal protein has at least 91% identity to any of the proteins depicted as SEQ ID Nos. 1 to 3. In a still further embodiment of the present invention the insecticidal protein has at least 92% identity to any of the proteins depicted as SEQ ID Nos. 1 to 3. In a still further embodiment of the present invention the insecticidal protein has at least 93% identity to any of the proteins depicted as SEQ ID Nos. 1 to 3. In a still further embodiment of the present invention the insecticidal protein has at least 94% identity to any of the proteins depicted as SEQ ID Nos. 1 to 3. In a still further embodiment of the present invention the insecticidal protein has at least 95% identity to any of the proteins depicted as SEQ ID Nos. 1 to 3. In a still further embodiment of the present invention the insecticidal protein has at least 96% identity to any of the proteins depicted as SEQ ID Nos. 1 to 3. In a still further embodiment of the present invention the insecticidal protein has at least 97% identity to any of the proteins depicted as SEQ ID Nos. 1 to 3. In a still further embodiment of the present invention the insecticidal protein has at least 98% identity to any of the proteins depicted as SEQ ID Nos. 1 to 3. In a still further embodiment of the present invention the insecticidal protein has at least 99% identity to any of the proteins depicted as SEQ ID Nos. 1 to 3. Preferably, the insecticidal protein according to the present invention comprises a motif depicted as -LPCCPG- and or -ICTPA- (SEQ ID Nos. 64 and 65 respectively). The percentage of sequence identity for proteins is determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the amino acid sequence in the comparison window may comprise additions or deletions

(i.e. gaps) as compared to the initial reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical amino acid residue occurs in both sequences to yield the number of match positions, dividing the number of match positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity. When calculating the percentage sequence identity the sequences may be aligned allowing for up to 3 gaps with the *proviso* that in respect of the gaps, a total of not more than 15 amino acid residues is affected. Optimal alignment of sequences for comparison may also be conducted by computerised implementations of known algorithms. In a particular embodiment of the present invention the sequence identity is calculated using the FASTA version 3 algorithm which uses the method of Pearson and Lipman (Lipman, D.J. and Pearson, W.R. (1985) Rapid and sensitive protein similarity searches and Science. 227:1435-1441 and Pearson, W.R. and Lipman, D.J. (1988) Improved tools for biological sequence comparison. PNAS. 85:2444-2448) to search for similarities between the reference sequence (also termed the query sequence) and any group of sequences (termed further sequences). Methods also exist in the art which enable the percentage sequence identity between polynucleotide sequences to be calculated.

The protein may differ from the basic insecticidal protein sequence (such as SEQ ID No. 1) by conservative or non-conservative amino acid substitutions. A conservative substitution is to be understood to mean that the amino acid is replaced with an amino acid with broadly similar chemical properties. In particular conservative substitutions may be made between amino acids within the following groups:

(i) Alanine and Glycine;

(ii) Serine and Threonine;

(ii) Glutamic acid and Aspartic acid;

(iii) Arginine and Lysine;

(iv) Asparagine and Glutamine;

(v) Isoleucine and Leucine,

(vi) Valine and Methionine;

(vii) Phenylalanine and Tryptophan.

In general, more conservative than non-conservative substitutions will be possible without destroying the insecticidal properties of the proteins. Suitable variant proteins in

accordance with the present invention may be determined by testing insecticidal properties of the protein using routine methods which are well known to the person skilled in the art. Such variant proteins may also be synthesised chemically using standard techniques.

The present invention still further provides an insecticidal protein as described above wherein the amino acid at position X_1 is modified. In a further embodiment of the present invention the amino acid at position X_1 is acetylated. In a still further embodiment of the present invention the amino acid at position X_1 is at the N-terminus. In a still further embodiment of the present invention the N-terminal region of the insecticidal protein comprises the sequence X_1X_2ICT- where X_1 and X_2 are any amino acid.

The present invention still further provides a polynucleotide encoding an insecticidal protein in its unmodified form.

The present invention still further provides a polynucleotide sequence which is the complement of one which hybridises to a polynucleotide as described above at a temperature of about 65°C in a solution containing 6 x SSC, 0.01% SDS and 0.25% skimmed milk powder, followed by rinsing at the same temperature in a solution containing 0.2 x SSC and 0.1% SDS wherein said polynucleotide sequence still encodes an insecticidal protein. In a further embodiment of the present invention the polynucleotide sequence comprises the sequence depicted as SEQ ID Nos. 4 to 14.

Further polynucleotide sequences according to the present invention may be identified from fungal DNA libraries. Suitable oligonucleotide probes may be constructed on the basis of the amino acid sequences of the proteins according to the present invention and used to screen any such DNA library for the identification of further polynucleotides encoding proteins according to the invention. In a still further embodiment of the present invention the amino acid sequences depicted as SEQ ID Nos. 1 to 3 may be used for the construction of oligonucleotide probes by the skilled man. In a still further embodiment of the present invention the sequences depicted as SEQ ID Nos. 4 to 14 may be used for the construction of oligonucleotide probes. In a still further embodiment of the present invention the DNA library is a *Paecilomyces spp.* DNA library. The person skilled in the art is well versed in methods for the production and screening of DNA libraries and the necessary techniques for the subsequent identification, isolation and sequence determination of polynucleotides which encode further insecticidal proteins in accordance with the present invention. The person skilled in the art will appreciate that alternative methods exist for the

identification and characterisation of related insecticidal sequences from other sources preferably other members of the *Paecilomyces* genus. Such methods include PCR strategies based on oligonucleotide primers using the sequence information provided herein or from sequences obtainable by the methods described above.

In a further aspect of the present invention there is provided an insecticidal synergistic combination comprising a first protein which is a protein as described above and at least one further protein. In a further embodiment of the present invention the further protein is an insecticidal CRY protein. The term "CRY protein" includes crystal endotoxin proteins (and secreted CRY) and the vegetative insecticidal proteins (and secreted VIP) which are active against insects including *Lepidoptera*, *Coleoptera* and *Diptera*. Such proteins are available *inter alia*, from the bacterium *Bacillus thuringienesis* and are well known to the person skilled in the art. Particularly preferred CRY proteins which may be used in accordance with the present invention include those proteins obtainable from *Bacillus thuringienesis* variety *tenebrionis* which has been deposited under the German Collection of micro-organisms (Deutsche Sammlung von Microorganism) under reference DSM 2803 or strains JHCC 4835 and JHCC 4353 deposited under the National Collections of Industrial and Marine Bacteria (Aberdeen) under the accession numbers NCIMB 40091 and 40090, respectively. In a still further embodiment of the present invention said further protein comprises a sequence selected from the group consisting of SEQ ID Nos. 54 to 59.

The present invention still further provides a polynucleotide which comprises regions encoding the first and further protein as described above. In a further embodiment of the present invention the polynucleotide comprises a region encoding a first protein which comprises the sequence depicted as SEQ ID No. 2. In a still further embodiment of the present invention the polynucleotide comprises a region comprising a sequence selected from the group depicted as SEQ ID Nos. 4 to 14. The insecticidal proteins or protein combinations according to the invention may be prepared in a number of ways which are apparent to the person skilled in the art. For example, by chemical synthesis using a standard peptide synthesiser, or using recombinant DNA technology to express the protein/combination in suitable organisms such as plants and micro-organisms such as *E. coli*, *Saccharomyces cerevisiae* or *Pichia pastoris*.

In a further aspect of the present invention there is provided a method of evolving a polynucleotide which encodes a protein having insecticidal properties comprising: (a)

providing a population of variants of said polynucleotide and further polynucleotides which encode further proteins, at least one of which is in cell free form; and (b) shuffling said variants and further polynucleotides to form recombinant polynucleotides; and (c) selecting or screening for recombinant polynucleotides which have evolved towards encoding a protein having the said insecticidal properties; and (d) repeating steps (b) and (c) with the recombinant polynucleotides according to step (c) until an evolved polynucleotide which encodes a protein having insecticidal properties has been acquired wherein said population of variants in part (a) contains at least a polynucleotide encoding a protein as described above. In a further embodiment of the present invention the evolved polynucleotide encodes an insecticidal protein having favourable properties for use in an applied context. For example enhanced activity or efficacy in a particular crop plant.

The present invention still further provides a method as described above wherein said population of variants in part (a) contains at least a polynucleotide encoding the protein depicted as SEQ ID Nos. 1 to 3 and said further polynucleotides in part (a) encode a CRY protein. The methods for evolving a polynucleotide as described above are well known to the person skilled in the art and are described *inter alia*, in US Patent No. 5,811,238.

The present invention still further provides a polynucleotide obtainable or obtained by the methods described above and a protein encoded by any such polynucleotide.

The present invention still further provides a DNA construct comprising in sequence a plant operable promoter operably linked to a polynucleotide encoding a protein as described above operably linked to a transcription termination region. In a further embodiment of the present invention the DNA construct further comprises a region or a plurality of regions which provide for the targeting of the protein product or products to a particular location or locations. For example, if it is desired to provide the protein outside of the cell then an extracellular target sequence may be ligated to the polynucleotide encoding the protein of the present invention. Other examples of targeting include targeting to a specific intracellular organelle or compartment such as a chloroplast, any other plastid, endoplasmic reticulum, peroxisome, the oil body, mitochondrion or vacuole. Numerous protein targeting sequences are available to the person skilled in the art and any of these sequences may be used to provide either (i) the protein according to the present invention *per se* and/or (ii) the further protein to, preferably, substantially the same location. In a still further embodiment of the present invention the target sequence comprises a sequence

selected from the group depicted as SEQ ID Nos. 15 to 19 or a polynucleotide encoding a protein selected from the group depicted as SEQ ID Nos. 20 to 24. The targeting polynucleotide sequence may be located 5' and/or 3' of the polynucleotide encoding the protein or combination according to the present invention.

The present invention still further provides a DNA construct as described above which further comprises a region which provides for the production of a protein which acts as a selectable marker. The selectable marker may, in particular, confer resistance to kanamycin; hygromycin or gentamycin. Further suitable selectable markers include genes which confer resistance to herbicides such as glyphosate based herbicides or resistance to toxins such as eutypine. Other forms of selection are also available such as hormone based selection systems such as the Multi Auto Transformation (MAT) system of Hiroyrasu Ebinuma *et al.* 1997. PNAS Vol. 94 pp2117-2121; visual selection systems which use the known green fluorescence protein, β glucoronidase and any other selection system such as mannose isomerase, xylose isomerase and 2-deoxyglucose (2-DOG).

The present invention still further provides a DNA construct as described above wherein the plant operable promoter is selected from the group consisting of *Agrobacterium rhizogenes* RolD; potato protease inhibitor II; CaMV35S; FMV35S; NOS; OCS; Patatin; E9; alcA/alcR switch; GST switch; RMS switch; oleosin; ribulose bisphosphate carboxylase-oxygenase small sub-unit promoter and other root specific promoters including MR7 promoter (maize); Gos 9 (rice) and GOS2 promoters. Terminators which can be used in the constructs according to the present invention include Nos, proteinase inhibitor II and the terminator of a gene of alpha-tubulin (EP-A 652,286). It is equally possible to use, in association with the promoter regulation sequence, other regulation sequences which are situated between the promoter and the sequence encoding the protein according to the present invention, such as transcriptional or translational enhancers, for example, tobacco etch virus (TEV) translation activator described in International Patent application, PCT publication number WO87/07644. The polynucleotide encoding the insecticidal protein or combination according to the invention may also be codon-optimised, or otherwise altered to enhance for example, transcription once it is incorporated into plant material. Examples of preferred codon usage from cotton and maize plants is set out in Table 1 below.

Table 1

Amino Acid	Cotton preference	Maize preference
Alanine	GCT	GCC
Arginine	AGG	AGG
Asparagine	AAC	ACC
Aspartic Acid	GAT	GAC
Cysteine	TGC	TGC
Glutamine	CAA	CAG
Glutamic Acid	GAG	GAG
Glycine	GGT	GGC
Histidine	CAT	CAC
Isoleucine	ATT	ATC
Leucine	CTT	CTG
Lysine	AAG	AAG
Methionine	ATG	ATG
Phenylalanine	TTC	TTC
Proline	CCT	CCG
Serine	TCT	AGC
Threonine	ACT	ACC
Tryptophan	TGG	TGG
Tyrosine	TAC	TAC
Valine	GTT	GTG

Such codon optimisation may also be used to alter the predicted secondary structure of the RNA transcript produced in any transformed cell, or to destroy cryptic RNA instability elements present in the unaltered transcript, thereby increasing the stability and/or availability of the transcript in the transformed cell (Abler and Green. 1996. Plant Molecular Biology (32) pp63-78). The expression of the protein and/or combination according to the present invention may also be enhanced through the inclusion of one or more intronic sequences within the polynucleotide encoding said protein and/or combination. (Rose and Beliakoff, 2000. Plant Physiology (122) pp.535-542). Examples of such sequences are the second

intron of the *Solanum tuberosum* LS1 gene and the alcohol dehydrogenase 1 gene (*adh1*) intron of monocotyledonous plant species. The chloroplast expression method (McBride et al. 1995. Biotechnology (13) pp362-365) may also be used to achieve enhanced expression of the protein and/or combination according to the present invention. This method is well known to the person skilled in the art and basically comprises transformation of the chloroplast genome with a polynucleotide under the control of a functional chloroplast-activated promoter or promoter/enhancer combination.

In a further aspect of the present invention there is provided a method of providing a plant or plant part with an insecticidal protein or an insecticidal protein synergistic combination comprising: (a) inserting into the genome of plant material a polynucleotide which encodes a protein as described above or a polynucleotide which comprises regions encoding the first and further protein as described above or a DNA construct as described above; or (a) inserting into the genome of plant material which is capable of producing a further protein, a polynucleotide encoding a first protein as described above; or (a) inserting into the genome of plant material which is capable of producing a first protein as described above, a polynucleotide which provides for a further protein; and (b) regenerating plants or plant parts from said material; and (c) selecting the plants or plant parts having said protein or combination. The polynucleotide/DNA construct may be incorporated into the cells by plant transformation techniques which are well known to the person skilled in the art. Such techniques include but are not limited to particle mediated biolistic transformation, *Agrobacterium*-mediated transformation, protoplast transformation (optionally in the presence of polyethylene glycols); sonication of plant tissues, cells or protoplasts in a medium comprising the polynucleotide or vector; micro-insertion of the polynucleotide or vector into totipotent plant material (optionally employing the known silicon carbide "whiskers" technique), electroporation and the like.

The present invention still further provides a method of providing a plant with an insecticidal protein synergistic combination comprising crossing a first plant which is capable of providing a first protein as described above with a second plant which is capable of producing a further protein and selecting the resultant plant which is capable of producing said combination.

The present invention still further provides plants or plant parts obtained according to the method as described above.

The present invention still further provides plants or plant parts as described above wherein said protein or the first protein of said combination is post translationally modified. In a further embodiment of the present invention said protein or the first protein of said combination is acetylated. In a still further embodiment of the present invention said protein or the first protein is modified/acetylated at the N-terminus. In a still further embodiment of the present invention the N-terminal region of the insecticidal protein/first protein comprises the sequence X_1X_2ICT- where X_1 and X_2 are any amino acid. In a further embodiment of the present invention X_1 and X_2 are selected from the group consisting of: Glycine; Lysine; Serine; Tyrosine; Alanine; Methionine; Threonine; Glutamic acid; Aspartic acid; Asparagine and Valine. In a still further embodiment of the present invention the amino acids at positions X_1 and X_2 are Serine and Tyrosine respectively. In a still further embodiment of the present invention the amino acid at position X_1 is Glutamine.

The present invention still further provides plants or plant parts as described above selected from the group consisting of melons, mangoes, soybean, cotton, tobacco, sugarbeet, oilseed rape, canola, flax, sunflower, potato, tomato, alfalfa, lettuce, maize, wheat, sorghum, rye, bananas, barley, oat, turf grass, forage grass, sugar cane, pea, field bean, rice, pine, poplar, apple, peaches, grape, strawberries, carrot, lettuce, cabbage, onion, citrus, cereal, nut plants or other horticultural crops. Plants and plant parts in accordance with the present invention show improved resistance or enhanced tolerance to an insect pest when compared to control-like or wild-type plants. Resistance may vary from a slight increase in tolerance to the pest to total resistance so that the plant is unaffected by the presence of pest (where the pest is severely inhibited or killed).

The present invention still further provides a method of providing a plant or plant part with a further desired agronomic trait comprising: (a) inserting into the genome of plant material a polynucleotide which provides for the desired agronomic trait; and (b) regenerating plants or plant parts from said material; and (c) selecting the plants or plant parts having said desired agronomic trait wherein said plant material is capable of producing an insecticidal protein or an insecticidal protein combination as described above; or crossing a first plant which plant is capable of producing an insecticidal protein or an insecticidal protein combination as described above with a second plant which provides for said further desired agronomic trait and selecting the resultant plant which is capable of producing the further agronomic trait. In a further embodiment of the present invention the said further

desired agronomic trait is selected from the group consisting of: herbicide resistance; insect resistance; nematode resistance; stress tolerance; altered yield; altered nutritional value or any other desirable agronomic trait. In a further embodiment of the present invention the further agronomic trait provides resistance to a herbicide which comprises glyphosate acid or agriculturally acceptable salt thereof.

The present invention still further provides plants or plant parts obtained according to the method of the preceding paragraph.

In a further aspect of the present invention there is provided an insecticidal protein comprising the sequence depicted as -X_1-X_2-X_3-Cys_4-X_5-X_6-X_7-X_8-X_9-X_{10}-Cys_{11}-X_{12}-X_{13}-X_{14}-X_{15}-X_{16}-Cys_{17}-Cys_{18}-X_{19}-X_{20}-X_{21}-X_{22}-Cys_{23}-X_{24}-X_{25}-X_{26}-X_{27}-X_{28}-X_{29}-X_{30}-X_{31}-Cys_{32}-X_{33}- (SEQ ID No. 60) wherein $X_{1\text{-}3, 5\text{-}10, 12\text{-}16, 19\text{-}22, 24\text{-}31}$ and $_{33}$ are any amino acid. In a further embodiment of the present invention the insecticidal protein comprises the sequence depicted as SEQ ID No. 60 and where X is any amino acid with the *proviso* that the amino acids a positions 14 and 15 are not cysteine. In a still further embodiment of the present invention the insecticidal protein comprises the sequence depicted as SEQ ID No. 60 where X is any amino acid other than cysteine. In the present case, the insecticidal peptides depicted as *inter alia*, SEQ ID Nos. 1 to 3 and 50 and the proteins encoded by SEQ ID Nos. 4 to 14 contain six cysteine residues all of which are believed to be involved in forming 3 intramolecular disulphide bonds. Thus the arrangement of the cysteine residues may be important in conferring insecticidal activity on the peptide. In a further embodiment of the present invention the amino acid at position X_1 is post translationally modified. In a still further embodiment of the present invention said amino acid at position X_1 is acetylated. In a still further embodiment of the present invention said amino acid at position X_1 is the N-terminus. In a still further embodiment of the present invention the N-terminal region of the insecticidal protein comprises the sequence X_1X_2ICT- where X_1 and X_2 are any amino acid. In a still further embodiment of the present invention X_1 and X_2 are selected from the group consisting of: Glycine; Lysine; Serine; Tyrosine; Alanine; Methionine; Threonine; Glutamic acid; Aspartic acid; Asparagine and Valine. In a still further embodiment of the present invention the amino acid at position X_1 is Glycine. In a still further embodiment of the present invention the amino acids at positions X_1 and X_2 are Serine and Tyrosine respectively. In a still further embodiment of the present invention the amino acid a position X_1 is Glutamine.

The present invention still further provides an insecticidal protein having a FASTA opt score greater than 109 when compared with SEQ ID No. 1 using FASTA Version 3. In a further embodiment of the present invention the insecticidal protein has a FASTA opt score greater than 110 when compared with SEQ ID No. 1 using FASTA Version 3. In a still further embodiment of the present invention the insecticidal protein has a FASTA opt score greater than 115 when compared with SEQ ID No. 1 using FASTA Version 3. In a still further embodiment of the present invention the insecticidal protein has a FASTA opt score greater than 117 when compared with SEQ ID No. 1 using FASTA Version 3. In a still further embodiment of the present invention the insecticidal protein has a FASTA opt score greater than 119 when compared with SEQ ID No. 1 using FASTA Version 3. In a still further embodiment of the present invention the insecticidal protein has a FASTA opt score greater than 120 when compared with SEQ ID No. 1 using FASTA Version 3. In a still further embodiment of the present invention the insecticidal protein has a FASTA opt score greater than 130 when compared with SEQ ID No. 1 using FASTA Version 3. In a still further embodiment of the present invention the insecticidal protein has a FASTA opt score greater than 140 when compared with SEQ ID No. 1 using FASTA Version 3. In a still further embodiment of the present invention the insecticidal protein has a FASTA opt score greater than 150 when compared with SEQ ID No. 1 using FASTA Version 3. The FASTA opt score may be calculated using the FASTA algorithm as described above. After computing the initial scores, FASTA determines the best segment of similarity between the query sequence and the further sequences using a the Smith-Waterman algorithm (Smith, T.F. and Waterman, M.S. (1981) Comparison of biosequences. Adv. Appl. Math. 2:482-489). The output is presented in the form of an opt score and this procedure is well known to the person skilled in the art.

The present invention still further provides an insecticidal protein obtainable or obtained from *Paecilomyces sp.* In a further embodiment of the present invention the insecticidal protein is obtained from *Paecilomyces farinosus*.

The present invention still further provides a method of controlling insects comprising providing at a locus where the insects feed, a protein or a protein combination as described above.

The present invention still further provides the use of a polynucleotide encoding an insecticidal protein as described above or a DNA construct as described above in a method

for the production of plants or plant parts which are resistant to insects. In a still further embodiment of the present invention the polynucleotide comprises the sequence selected from the group depicted as SEQ ID Nos. 4 to 14.

The present invention still further provides the use of a protein a or a protein combination as described above as an active ingredient of a pesticide.

The present invention still further provides the use of a *Paecilomyces Sp.* in the preparation of a pesticide containing as an active ingredient, a protein as described above. In a further embodiment of the present invention said *Paecilomyces Sp.* has been modified to allow for increased production of a protein as described above. The person skilled in the art can modify said *Paecilomyces Sp.* so that it is capable of producing the insecticidal proteins at levels which are increased compared to unmodified control-like *Paecilomyces Sp.* using techniques well known within the art. For example, by modifying the promoter elements attached to polynucleotide encoding the insecticidal protein as described above to increase protein production. In addition to this or alternatively, by inserting a second copy of the gene encoding the protein as described above into the selected *Paecilomyces Sp.* The production of the insecticidal protein according to the present invention may also be increase through standard strain improvement programs known to those skilled in the art.

The present invention still further provides a recombinant micro-organism which provides for production of a protein or a protein combination as described above. In a further embodiment of the present invention the micro-organism is an endophyte. An endophyte is generally accepted within the art as a micro-organism having the ability to enter into non-pathogenic endosymbiotic relationships with a plant host. A method of endophyte-enhanced protection of plants has been described in a series of patent applications by Crop Genetics International Corporation (for example, International Application Publication Number WO90/13224, European Patent Publication Number EP-125468-B1, International Application Publication Number WO91/10363, International Application Publication Number WO87/03303). International Patent Application Publication Number WO94/16076 (ZENECA Limited) describes the use of endophytes which have been genetically modified to express a plant-derived insecticidal peptide.

The present invention still further provides a recombinant baculovirus which comprises a protein or a protein combination as described above. The present invention still

further provides the use of a baculovirus according to the preceding sentence in a method of controlling insects.

According to a further aspect of the present invention there is provided an insecticidal protein which is capable of reacting with a monoclonal antibody raised to the protein selected from the group depicted as SEQ ID No. 1 to 3. The present invention still further provides an insecticidal protein which is capable of reacting with a polyclonal antibody raised to the protein selected from the group depicted as SEQ ID No. 1 to 3. Such antibodies may be generated and used to identify other proteins within the ambit of the present invention according to well known techniques within the art.

The present invention still further provides a composition comprising an insecticidally effective amount of a protein or a protein combination as described above and optionally an agriculturally acceptable carrier and/or a diluent and/or an insect attractant. The composition may be applied to the insects or to the environment in which they live, in particular, to plant parts or the surrounding soil, using standard agricultural techniques for example spraying. The insecticidal proteins and combinations according to the present invention may also be combined in application with other agrochemicals such as herbicides, fungicides and other insecticidal compounds including other insecticidal proteins. Examples of possible mixture partners include insecticidal lectins, insecticidal protease inhibitors and insecticidal proteins derived from species of the *Bacillus thurigiensis*, *Xenorhadus nematophilus*, or *Photorabdus luminescens* and other chemicals for example pyrethroids, carbamates, imidacloprid, organochlorines, macromolecules such as spinosad abamectin or emamectin.

The present invention still further provides a polynucleotide having a first region encoding a protein as described above and a second region encoding a further protein. The regions may be separated by a region which provides for a self processing polypeptide which is capable of separating the proteins such as the self processing polypeptide described in US5,846,767 or any similarly functioning element. Alternatively the regions may be separated by a sequence which acts as a target site for an external element which is capable of separating the protein sequences. Alternatively the polynucleotide may provide for a polyprotein which comprises a plurality of protein functions. In a further embodiment of the present invention the proteins of the polyprotein may be arranged in tandem. In a still further embodiment of the present invention the polyprotein comprises a plurality of protein

functions which are separated by linker sequences. Such polyproteins may comprise the proteins and/or further proteins according to the present invention and optionally further proteins such as those encoding any desired agronomic trait.

The present invention still further provides a plant cell comprising a protein or protein combination as described above or a polynucleotide encoding an insecticidal protein and/or an insecticidal protein combination as described above.

The present invention still further provides an insecticidal protein comprising the motif depicted as -LPCCPG- (SEQ ID No 63) and/or -ICTPA- (SEQ ID No. 64).

The insects to be controlled by the proteins of the present invention include the plant chewing insects and the plant chewing stages of insects such as insect larvae including: *Coleoptera*, *Lepidoptera*, *Orthoptera* and *Drosophila*, including, but not limited to: *Acanthoscelides obtectus*, *Bruchus sps.*, *Callosobruchus sps.* (bruchid beetles), *Agriotes sps.* (wireworms), *Amphimallon sps.* (chafer beetles), *Anthonomus grandis* (cotton boll weevil), *Ceutorhynchus assimilis* (cabbage seed weevil), *Cylas sps.* (sweet potato weevils), *Diabrotica sps.* (corn root worms), *Epicauta sps.* (black blister beetles), *Epilachna sps.* (melon beetles etc.), *Leptinotarsa decemlineata* (Colorado potato beetle) *Meligisthes sps.* (blossom beetles), *Melolontha sps.* (cockchafers), *Phyleotreta sps.*, *Psylliodes sps.* (flea beetles), *Popillia japonica* (Japanese beetle), *Scolytus sps.* (bark beetles), *Sitophilus sps.* (grain weevils), *Tenebrio molitor* (yellow mealworm), *Tribolium sps.* (flour beetles), *Trogoderma granarium* (Khapra beetle), *Acleris sps.* (fruit tree tortrixs), *Acraea acerata* (sweet potato butterfly), *Agrotis sps.* (cutworms), *Autographa gamma* (silver-Y moth), *Chilo sps.* (stalk borers), *Cydia pomonella* (codling moth), *Diparopsis sps.* (red bollworms), *Ephestia sps.* (warehouse moths), *Heliothis sps.*, *Helicoverpa sps.* (budworms, bollworms), *Mamestra brassicae* (cabbage moth), *Manduca sps.* (hornworms), *Maruca testulalis* (mung moth), *Mythimna sps.* (cereal armyworms), *Ostrinia nubilalis* (European corn borer), *Pectinophora gossypiella* (pink bollworm), *Phthorimaea operculella* (potato tuber moth), *Pieris brassicae* (large white butterfly), *Pieris rapae* (small white butterfly), *Plodia interpunctella* (Indian grain moth), *Plutella xylostella* (diamond-back moth), *Sitatroga cerealella* (Angoumois grain moth), *Spodoptera sps.* (armyworms), *Trichoplusia ni* (cabbage semilooper), *Acheta sps.* (field crickets), *Gryllotalph sps.* (mole crickets), *Locusta migratoria* (migratory locust), *Schistocerca gregaria* (desert locust), *Acrythosiphon pisum* and *Drosophila sp.*

The invention will now be described by way of the following non-limiting examples in combination with the following figures and sequence listing of which:

FIGURE 1 - is a schematic diagram of the organisation of the *Paecilomyces farinosus* gene.

FIGURE 2 - shows illustrates the signal-gene fusion part of a construct suitable for the transformation of corn.

FIGURE 3 - shows a construct suitable for the transformation of corn and illustrates the backbone vector pat UB1 poly2.

FIGURE 4 - Shows genomic map of the gene isolatable from *Paecilomyces farinosus*.

FIGURE 5 - Shows the vector pCR2.1 TOPO.

SEQ ID Nos. 1 to 3 - Insecticidal proteins obtainable from *Paecilomyces spp.* Also referred to as "R524445 protein" and "445 protein".

SEQ ID Nos. 4 to 6 - Polynucleotides encoding the insecticidal proteins.

SEQ ID Nos. 7 and 8 - Polynucleotides encoding the insecticidal proteins - codon optimised. For SEQ ID No. 7 - signal peptide is present from position 1 to 72 and the mature protein encoding sequences is from 73 to 174 (including the stop). For SEQ ID No. 8 - signal peptide is present from position 1 to 72 and the mature protein encoding sequences is from 73 to 174 (including the stop).

SEQ ID Nos. 9 and 10 - Polynucleotides encoding the insecticidal proteins - containing intron sequences. For SEQ ID No. 9 - signal peptide is present from position 1 to 68 and the intron sequence is from 99 to 288. SEQ ID No. 10 has two additional amino acids Serine and Tyrosine at the N-terminus of the mature protein. Signal peptide is present from position 1 to 72, Ser and Tyr are encoded by nucleotides 73 to 78 and the intron sequence is from 106 to 294.

SEQ ID No. 11 - Polynucleotide encoding the insecticidal proteins with two amino acids substituted for Serine and Tyrosine at the N-terminus of the mature protein. Signal peptide is present from position 1 to 72. Ser and Tyr are encoded by nucleotides 73 to 78 and the intron sequence is from nucleotides 100 to 288.

SEQ ID No. 12 - Polynucleotide encoding the insecticidal proteins containing intron and codon optimised. Signal peptide is present from position 1 to 78 and the intron sequence is from 105 to 288.

SEQ ID No. 13 - Genomic sequence of insecticidal protein obtainable from *Paecilomyces spp.* Signal peptide is present from position 57 to 107. The mature protein encoding sequences is from 108 to 397 (including the stop).

SEQ ID No. 14 - Polynucleotide encoding the mature insecticidal protein.

SEQ ID No. 15 to 19 - Polynucleotide sequences encoding the signal peptides from Dahlia (Dm-AMP-1), Radish (Rs-AFP1), Maize (hydroxyproline-rich glycoproten (HRGP)), Tobacco (PR-1a signal) and *Paecilomyces* respectively.

SEQ ID Nos. 20 to 24 - Protein sequences for the signal peptides from Dahlia (Dm-AMP-1), Radish (Rs-AFP1), Maize (hydroxyproline-rich glycoproten (HRGP)), Tobacco (PR-1a signal) and *Paecilomyces* respectively.

SEQ ID Nos. 25 to 53 - Primers.

SEQ ID Nos. 54 to 59 - Protein sequences for insecticidal proteins cry1Ia1 (Embl. Accession No. X62821); cry1Ia2 (Embl. Accession No. M98544); cry1Ia3 (Embl. Accession No. L36338); cry1Ia4 (Embl. Accession No. L49391); cry1Ia5 (Embl. Accession No. Y08920) and cry1Ib1 (Embl. Accession No. U07642) respectively.

SEQ ID No. 60 - Insecticidal protein sequence having cysteine residues in particular positions.

SEQ ID No. 61 - Polynucleotide encoding the genomic insecticidal protein sequence. Signal peptide is present from position 57 to 107. The mature protein encoding sequences is from 108 to 397 (including the stop).

SEQ ID No. 62 - Polynucleotide encoding insecticidal protein obtainable from *Paecilomyces sp.* Signal peptide is present from position 110 to 160. The mature protein encoding sequences is from 161 to 262 (including the stop).

SEQ ID No. 63 and 64 - Protein motifs.

SEQ ID No. 65 - Protein region.

EXAMPLES

Example 1

Culturing of *Paecilomyces farinosus*

Paecilomyces farinosus was routinely cultured on potato dextrose agar plates. Spores were harvested from the plates by adding sterile water and scraping with a sterile spatula. For production of insecticidal peptide $6x10^7$ spores were inoculated into 5x 200ml of SDB

medium in 500ml flasks. Cultures were incubated at 24°C with shaking at 180rpm for 7 days before harvest.

Example 2

Purification of insecticidal peptide

500ml of 7d culture filtrate was filtered through Whatman™ GF/B paper to remove mycelium and the supernatant diluted 4 fold in 20mM MES pH6. The supernatant was then loaded onto a S-Sepharose FF XK16/10 column (Pharmacia Biotech) previously equilibrated with 20mM MES pH6. Unbound protein was washed through the column with 3 column volumes of 20mM MES pH6 and bound protein was eluted with a linear gradient of 0-1M NaCl in 20mM MES pH6 over 20 column volumes. The eluate was monitored for peptide by online measurement of absorbance at 280 and 210nm. 5ml fractions were collected and following dialysis against 50mM Sodium Phosphate buffer pH7 assayed against *Heliothis virescens*.

Active fractions eluted around 250mM NaCl. These fractions were pooled and following concentration on Polyethylene glycol Mwt 20,000, were further purified by reverse phase. 2ml of sample was loaded onto a 3ml Resource RPC column (Pharmacia Biotech) and bound peptide eluted with a linear gradient of 0.05% trifluoroacetic acid (TFA) to 50% acetonitrile, 0.0% TFA over 20 column volumes. The eluate was monitored for absorbance at 210 and 280nm. The active peak eluted at approximately 20% acetonitrile.

Example 3

Identification of peptide sequence

The sequence of the active peptide in the product of Example 2 could not be determined directly, probably due to a blocked N-terminus. The peptide was reduced and subjected to and tryptic digestion. This yielded a series of fragments which could be sequenced using Edman degradation methods. Using a combination of this, and mass spectrometry, the sequence of the peptide was determined as being SEQ ID No. 2. The mass spectrometry data indicates that the N-terminal glycine is acetylated.

Example 4

Biological Activity in insect bioassay

The isolated peptide was bioassayed against a range of insect species using the following method:

Prior to the assay twenty neonate *lepidoptera* larvae were gently brushed into each of three 'minipots' containers per treatment (i.e. three replicates per treatment). The peptide from Example 2 was diluted using 0.1% Synperonic™ solution to act as a wetter and aid the spread of the material over the waxy leaf cuticle. In spectrum assays, test materials were made up to a single high concentration, whereas in potency assays vs. *H. virescens* a rate range was tested.

Three freshly excised cotton leaves per treatment had 0.1 ml of the appropriate treatment applied by pipette to the centre of the axial surface of each leaf. The droplet was then spread over a circular area in excess of the diameter of a minipot with a fine artists paint brush (a fresh paint brush being used for each compound to avoid contamination). The leaves were left in a fume cupboard just long enough for the surface deposit to dry but care was taken to avoid excessive leaf wilting.

Once dry the leaves were placed, contaminated surface down over the appropriately labelled minipot and a lid snapped over it. The minipots were placed in plastic trays and held in a controlled temperature at 25-27°C.

After three days the numbers of live larvae remaining were counted and percent mortality determined. In the *H. virescens* potency assay the test data was run through a logit analysis package to establish the LC_{50}.

The results for 4 lepidopteran pests are shown in Table 2.

Table 2

Test Species	Rate (ppm)	% kill
Heliothis virescens	1000	100
Helicoverpa zea	1000	100
Spodoptera exigua	1000	100
Plutella xylostella	1000	100

b) Cell cytotoxicity

Two cell lines were used to determine if the peptide from Example 2 was cytotoxic to either mammalian cells (MEL cells) or insect cells (Sf21 cells). MEL cells and Sf21 cells were grown in DMEM and TC100 media respectively in 96-well microtitre plates and incubated with the appropriate concentration of peptide. The cells were scored for visible cell death after 24 hours and viability and growth assessed after 3 (MEL cells) or 4 (Sf21) days using the reduction of MTT (3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) to form an insoluble purple formazan as a marker for metabolically active cells. At the highest rate tested (100 µg/ml) the peptide did not inhibit cell growth or cause any cytotoxic effects on either cell line.

Example 5

Comparison of protein sequences to SEQ ID N0. 1 using the FASTA Algorithm:

A FASTA comparative search of SEQ ID No. 1 to a database of protein sequences was carried out. SEQ ID NO 1 was compared to all publicly available protein sequences using the FASTA method (FASTA version 3.0t82 November 1, 1997 Reference: W.R. Pearson & D.J. Lipman PNAS (1988) 85:2444-2448). The results were given in the form of an opt score.

Example 6

Characterisation of natural coding sequence of peptide of SEQ ID No. 2

Harvesting of Material

A *Paecilomyces farinosus* strain having insecticidal activity was grown in Sabouraud Dextrose Broth (Difco Laboratories: 10g Bacto Neopeptone, 20g Bacto Dextrose per litre water) for 5 days at 24°C with shaking at 180 rpm. The culture was pelleted (8000rpm, 10 minutes) and stored at -80°C until use.

RNA Extraction

Harvested material was ground to a fine powder using a pestle and mortar under liquid nitrogen. RNA was extracted from 1g of fungal pellet using the Qiagen RNeasy kit,

following manufacturers specifications. The total RNA fraction was eluted from the RNeasy purification column in 1ml water. Poly(A)+ RNA was isolated from 700µg total RNA using the Promega PolyATtract mRNA isolation system I, following manufacturers' specifications. The Poly(A)+ RNA fraction was eluted from the magnetic beads in 1ml water, and concentrated to 15µl (of approximately 0.5mg/ml) by ethanol precipitation.

RNA samples were stored at -80°C until use.

In the following reactions, the primers and probes used are summarised in Table 3.

Table 3

Primer/Probe sequences

Designation	Primer Sequence	SEQ ID No
Anchor1	TCGGGCTCGCATGAATTCGCGGCCGCATTTTT TTTTTTTTTTT	25
Anchor1 - R1	TCGGGCTCGCATGAATTCG	26
Anchor1 - R2	ATGAATTCGCGGCCGCAT	27
Anchor1 - R3	TCGGGCTCGCATGAATTCGCG	28
Anchor1 - R4	CTCGCATGAATTCGCGGCCGC	29
F1	ATHTGYACNCCNGCNGG	30
F2	ATHTGYACNCCNGCNGGNGT	31
F3	ACNCCNGCNGGNGTNAA	32
F4	CCNTGYTGYCCNGGNYT	33
F5	TNAARTGYATHGGNGG	34
F6	GGNGTNAAYAAYAARGTNTG	35
F7	AARATHTGYACICCIGCIGGIGTIAA	36
F8	CCIGCIGGIGTIAARTGYCCIGCIGC	37
F9	TGYCCIGCIGCIYTICCITGYTGYCC	38
F10	TGYATHGGIGGIGTIAAYAAYAARGT	39
F11	TAAATGTCCCGCGGCTCTTCC	40
F12	CGGCTCTTCCTTGCTGCCCCG	41
F13	TGCTGCCCCGGACTTCGCTGC	42
Anchor3- attach	HO-GGTTTAATTACCCAAGTTTGAGNNNNN - NH_2	43

Anchor3	PO_4 - CTCAAACTTGGGTAATTAAACC - NH_2	44
Anchor3 - F1	GGTTTAATTACCCAAGTT	45
Anchor3 - F2	TAATTACCCAAGTTTGAG	46
Anchor3 - F3	GGTTTAATTACCCAAGTTTGAG	47
R1	CAIACYTTRTTRTTIACICCICC	48
R2	ATGCAGCGAAGTCCGGGGCAG	49
R3	GGGGCAGCAAGGAAGAGCCGC	50
R4	AAGAGCCGCGGGACATTTAAC	51
Probe F	AGTTAAATGTCCCGCGGCTCTTCCTTGCTGCC CCGGACTTCGCTGCATC	52
Probe R	GATGCAGCGAAGTCCGGG	53

where "F" designates a forward primer and "R" designates a reverse primer.

RACE PCR

First strand cDNA synthesis

The ClonTech Advantage RT-for-PCR kit was used for this, in accordance with the manufacturers' specifications.

20 pmol oligo(d)T primer 'Anchor1' was annealed to 1 µg total RNA in a total volume of 13.5 µl by heating to 70°C for 2 minutes and rapidly quenching on ice.

The following reaction components were added:

4 µl 5 x reaction buffer; 0.5µl RNase inhibitor; 1 µl MMLV reverse transcriptase; 1 µl 10mM dNTP. The reaction was incubated at 42°C for 1 hour, and the reaction stopped by denaturing the enzyme at 95°C for 5 minutes. 80 µl RNase free water was added and cDNA stored at -80°C.

3' RACE PCR

5 µl of reaction mix from the first strand cDNA synthesis reaction was used as a template with various primer set combinations to amplify the 3' end of the peptide coding cDNA. The primers (see Table 3) used were degenerate, and designed on the known amino acid sequence of the N-terminal end of the mature peptide to allow for selective amplification.

PCR reactions were performed using Ready-To-Go PCR Beads (Amersham Pharmacia Biotech). These contain all necessary components for a PCR reaction as a bead in a 0.5 ml tube. The following components were added: cDNA template 5 µl reaction mix from cDNA synthesis step; Forward primer 25 pmol; Reverse primer 25 pmol; Sterile Water to a final volume of 25 µl .

PCR cycle conditions

(1) 95 °C 1 min; (2) 95 °C 1 min; (3) 64 °C* 1 min; (4) 72 °C 1 min (steps 2-4 for 30 cycles); (5) 72 °C 10 min (*Annealing temperature varied depending on primer set).

PCR products were visualised by agarose gel electrophoresis on a 1% agarose gel in TBE buffer. Discrete PCR products were cloned into pCR2.1 TOPO using the Invitrogen TOPO TA cloning kit according to the manufacturers' specification. Each ligation contained: 1 µl PCR product; 1 µl pCR2.1 TOPO vector (Figure 5); 3 µl Sterile Water and was incubated at room temperature for 5 minutes. 2 µl of each ligation mix was transformed into TOP10 competent cells by heat shock at 42°C for 30 seconds followed by incubation on ice for 2 minutes. Transformed cells were allowed to express beta-lactamase by incubation at 37°C in SOC medium (2% tryptone, 0.5% yeast extract, 10mM NaCl, 2.5mM KCl, 10mM $MgCl_2$, 10mM $MgSO_4$, 20mM glucose) for 1 hour with shaking at 225 rpm.

Cells were plated on Luria-Bertani Agar plates (1.0% tryptone, 0.5% yeast extract, 1.0% NaCl, 15g/L agar, 0.006% X-gal, 0.15mM IPTG) containing 50µg/ml kanamycin for plasmid transformant selection and to enable identification of those containing recombinant TOPO TA isolates. Discrete white colonies were selected from different PCR TOPO TA reactions, grown overnight in 5 ml Luria-Bertani (1.0% tryptone, 0.5% yeast extract, 1.0% NaCl, in water, pH 7.0) containing 50µg/ml kanamycin.

Plasmid DNA was extracted from the cultures using the Wizard DNA purification kit (Promega), following manufacturers' specifications. DNA was eluted in 50µl sterile water. Plasmid DNA was digested with *EcoRI* to confirm the presence and size of inserts.
3 µl Plasmid DNA; 1 µl *EcoRI* (Kramel Biotech); 1 µl 10 x Restriction Buffer 6 (Kramel Biotech); 5 µl Sterile water. Digests were incubated at 37°C for 2 hours and the presence or absence and size of inserts determined by agarose gel electrophoresis. Based on these analyses, recombinant plasmids were selected for sequencing on a Perkin Elmer ABI 377XL DNA sequencer with the ABI Prism dye terminator cycle sequencing ready reaction kit,

according to the manufacturers' protocol. 4 pmol primer M13 Univ or M13R; 5 µl DNA; Sterile water to 12 µl. The coding sequence of the peptide of SEQ ID No. 2 was identifiable by translation of the nucleotide sequence into amino acid sequence in all possible reading frames and comparison of this sequence to the known amino acid sequence of the peptide. This analysis used the DNA Star sequence analysis software (SeqMan, EditSeq, Macaw, VectorNTI).

5' RACE PCR

An anchor-ligation approach (Troutt, A.B., et al., *Proc. Natl. Acad. Sci. USA.* 89, 9823-9825) was used to obtain the nucleotide sequence of the 5' half of the 524445 gene. This entailed attachment of a specific anchor primer to the 5' end of the first strand cDNAs. Use of this sequence together with mRNA specific for the peptide of SEQ ID No. 2 complementary 3' primers allowed for selective amplification of the 5' end of the corresponding coding cDNA.

Primer annealing

Complementary oligonucleotides Anchor3 and Anchor3-attachment were annealed to each other in equimolar ratio's at three different final concentrations (1nM, 100nM, 10mM). Oligonucleotide mixtures were heated to 95°C and cooled slowly to 45°C for annealing.

(2) Ligation of annealed primer to cDNA

The attachment primer is complementary to the anchor primer, but contains a 3' extension of 5 additional fully degenerate bases i.e. synthesised with A, G, C and T at each position. This degenerate 'tail' allows individual attachment primers to anneal to the 3' terminus of any cDNA molecule. An amido group at the 3' end of the primer blocks DNA synthesis. A phosphate group at the 5' end of the anchor primer allows ligation of this to the 3' end of the cDNA molecules to provide a specific recognition sequence for PCR amplification. Reactions for ligation of annealed anchor primers to first strand cDNA preparations contained: 5 µl reaction mix from first strand cDNA synthesis; 30 mM Tris HCl (pH 8); 10 mM $MgCl_2$; 10 mM Dithiothreitol; 0.5 mM ATP; 1 µl T4 DNA ligase (4 U/µl) (Kramel Biotech); 1 µl Water; 1 µl annealed anchor primers (final concentrations of 100mM, 10nM, 1mM).

Reactions were cycled overnight as follows:

25 °C 5 min

↕ Slow Ramp Rate

4 °C 5 min

Reactions were pooled, incubated at 95°C for 5 minutes and snap frozen in liquid nitrogen. After thawing on ice, excess primers were removed by purification through a Wizard PCR clean-up column (Promega) using the manufacturers' specifications. cDNAs were eluted in 40 µl water.

(3) RACE PCRs

PCR reactions were set up using the anchor-linked cDNA as a template, specific forward primers based on this anchor sequence, and specific primers based on the gene sequence of the peptide of SEQ ID NO 1 identified previously by 3'RACE. PCR reactions were performed using Ready-To-Go PCR Beads (Amersham Pharmacia Biotech) as used previously for 3' RACE. Components added to the PCR beads were: 1 µl cDNA template with Anchor3 annealed to 3' end of first strand cDNA; 20 pmol forward primer; 20 pmol reverse primer; sterile water to total volume of 25 µl.

PCR cycle conditions were: (1) 95 °C 1 min; (2) 95 °C 1 min; (3) 58 °C* 1 min; (4) 72 °C 1 min (steps 2-4 for 30 cycles); (5) 72 °C 10 min (*Annealing temperature varied depending on primer set). PCR products were visualised by agarose gel electrophoresis and TOPO cloned as described above. Plasmid DNA was extracted from clones carrying candidate recombinant plasmids by Wizard miniprep, *EcoRI* digested and sequenced, as performed previously for 3' RACE clones (described above).

cDNA Library

Library Construction

A cDNA library of the fungus *Paecilomyces farinosus* was constructed using the lambda-ZAP cDNA synthesis and ZAP-cDNA Gigapack III Gold Cloning kit from Stratagene, according to the manufacturers' specifications unless stated. Double stranded cDNA was synthesised using 5 µg the mRNA from the peptide of SEQ ID No. 2 (see above) as a template. This involved first and second strand cDNA synthesis, blunting of cDNA termini, ligation of adapters, and digestion with specific restriction enzymes to produce

appropriate 'sticky ends' for directional cloning. A Sephacryl S-400 HR MicroSpin column (Amersham Pharmacia Biotech) was used to remove excess adapters rather than the size fractionation step suggested in the kit. The gel filtration medium provided in the kit (sepharose CL-2B) separates molecules on the basis of size with a cut-off of 400bp. As the mature insecticidal peptide is only 33 amino acids long, it is highly likely that the gene may be smaller than 400bp and would have been selected against using the sepharose filtration medium. cDNAs were ligated into the Uni-ZAP XR vector and packaged into phage. The library titre was 2.5 million clones, with an average insert size of 700bp ranging from 150bp to 2 Kb.

Library screening

A total of 500,000 plaques were plated on Luria-Bertani Agar plates according to the cDNA library manufacturers' specification. Duplicate lifts of the plaques were made onto nitrocellulose membrane (Hybond-N, Amersham Pharmacia Biotech).

The membranes were prehybridised in Denhardts hybridisation solution (5x SSPE, 5x Denhardt's Reagent [50x Denhardt's Reagent: 5g Ficoll, 5g polyvinylpyrrolidone, 5g bovine serum albumin, sterile water to 500ml], 0.5% SDS, sterile water to 1L) containing 200µl salmon sperm DNA (10mg/ml) which had been denatured by boiling for 10 minutes, for 2 hours at 65°C. A radioactive probe was prepared by end labelling an oligonucleotide specific for the coding sequence of a peptide of SEQ ID NO 1: 25 ng Oligonucleotide (445-F11); 1 µl polynucleotide Kinase Buffer (Kramel Biotech); 1.5µl T4 Polynucleotide Kinase (Kramel Biotech); 5 µl gamma 32P dATP; sterile Water to 10 µl. The probe was incubated at 37°C for 5 hours. The probe was added to 50ml Denhardts hybridisation solution and hybridised overnight at 65°C. Membranes were washed in a 0.1x SSC, 0.1% SDS solution for 4 x 15 minutes to remove unbound probe. Exposure to x-ray film identified positive plaques containing sequence encoding protein depicted as SEQ ID No. 2.

Positive plaques were cored from the original agar plates into 1ml SM buffer (5.8g NaCl, 2g $MgSO_4.7H_2O$, 50ml 1M Tris-Hcl pH 7.5, 5ml 2% gelatin, sterile water to 1L) containing 20ml chloroform and vortexed. The phage DNA was allowed to enter the phage buffer by incubation at 4°C overnight. Samples of phage were then diluted and re-plated to obtain approximately 200 plaques per plate. The plaque lift and hybridisation procedure above was repeated to identify positives.

This process was followed for three rounds of screening until the plaques were pure. Self excision of 12 candidate positive plaques into colonies was performed as per Stratagene's specifications with the cDNA library kit.

Candidate colonies were grown overnight in 5ml Luria-Bertani medium containing 100µg /ml ampicillin, plasmid DNA extracted using Promega's Wizard miniprep kit, and the inserts sequenced using M13 Universal and M13 Reverse primers (see above for details of all).

Nucleotide Sequence

The nucleotide sequence of the peptide of SEQ ID No. 2 is shown as SEQ ID No. 14 and also in Figure 2. The putative translation initiation codon and stop codon are shown in italics. The sequence which codes for the mature peptide is underlined.

The sequence in Figure 2 indicates that there is approximately a 110 nucleotide 5' non-coding sequence and a 160 nucleotide 3' non-coding sequence. There seems to be a 17 amino acid signal peptide 5' of the mature coding sequence. Potential signal sequence cleavage sites were predicted based on the method of von Heijne, G. (1986). *Nucleic Acids Research.* **14**, 4683. The potential cleavage site is indicated by a downward pointing arrow. It is probable that a secondary processing event removes the signal peptide from the mature peptide, e.g. by signal peptidase cleavage.

Example 7

Expression of peptide in corn

European Corn Borer (*Ostrinia nubilalis*) and therefore corn, *Zea mays*, which has been transformed so as to express this peptide would be expected to be protected against this pest.

Suitable constructs for expression in corn can be summarised as follows:

Construct	*Promoter*	*Signal Peptide*	*Gene*	*Terminator*
1	Maize Ubi	SEQ ID NO 17	SEQ ID No. 14^{+}	nos
2	Maize Ubi	SEQ ID NO 15	SEQ ID No. 14^{+}	nos
3	Maize Ubi	SEQ ID NO 19*	SEQ ID No. 14^{+}	nos

* This signal peptide contains an internal *NcoI* site which can be mutated (for example CCATGG → CTATGG) to destroy it if *NcoI* is required for cloning.

+ The natural coding sequence can be modified in accordance with the degeneracy of the genetic code, and in particular for the purpose of codon optimisation in corn.

The signal peptide can be fused to the mature gene for example using an overlapping PCR approach as illustrated in Figure 2. The fusion is suitably designed with restriction sites to allow cloning into monocot vectors. For example, it may comprise the following:

5' *NcoI - KpnI* -------- Signal -------- Gene -------- *SacI - HindIII* 3'

The full length signal-gene fusion can be ligated between the maize ubiquitin promoter and nos terminator into a backbone vector containing PAT selection (phosphinothricin - basta herbicide resistance).

These constructs can be used to transform corn cells which can then be grown into callus as is well known in the art. The transformed callus can be subjected to a corn callus transient assay and/or an *in vivo* bioassay to confirm expression and activity of the peptide.

Example 8

Expression of Peptide in Cotton

The peptide of the invention has good activity against the Beet Armyworm (*Spodoptera exigua*) which is a major cotton pest. Thus cotton *Gossypium hirsutum*, which has been transformed to express this peptide would be protected against this pest.

Suitable constructs for use in the transformation in this case can be summarised as:

Construct	*Promoter*	*Signal Peptide*	*Gene*	*Terminator*
4	RolDFd	SEQ ID NO 23	SEQ ID NO 14	potato protease inhibitor II
5	RolDFd	SEQ ID NO 20	SEQ ID NO 14	potato protease inhibitor II
6	RolDFd	SEQ ID NO 24	SEQ ID NO 14	potato protease inhibitor II

The signal peptide can be fused to the mature gene using an overlapping PCR approach as in Example 7. In this case, the fusion is suitably designed with restriction sites to allow cloning into dicot vectors. The full length signal-gene fusion can be ligated into a housekeeping vector between the RolDFd promoter and potato protease inhibitor II terminator. The entire cassette could then be cut out using restriction enzymes and ligated

into an appropriate binary vector. Constructs can then be tested using conventional methods. Other modifications of the present invention will be apparent to those skilled in the art without departing from the scope of the invention.

Example 9

Insecticidal activity of the protein combination

Previously prepared European Corn Borer (ECB) artificial diet was dispensed in small quantities into tubes and held in a warm water bath at 70°C. To each tube containing 975*ml* of diet, 75*ml* of the appropriate test sample was added. The test samples comprised a mixture of the cry1Ia1 protein (SEQ ID No. 54) and the protein depicted as SEQ ID No. 2 (the R524445 protein). The "incorporated diet" was mixed well and 180*ml* aliquots were then pipetted out onto Falcon™ 1006 petri dishes, giving five replicates for each sample. The dishes were infested 1 - 5 hours after the diet is dispensed with five 1st instar larvae per dish/rep and then lidded. The test was held in the dark at 27°C and 70 - 80% RH and the insects were assessed five days after treatment for mortality. The results are shown in the Table 4 below:

Table 4

(Results shown as number of insects living)/(number of insects within the dishes)						
cry1Ia1 Conc. (PPM)	R524445 Protein concentration (PPM)					
	0.00	4.91	7.60	11.75	18.16	22.4
0.00	15/15	13/15	10/14	8/15	6/15	5/15
2.48	10/11	1/14	0/15	0/15	0/15	-
3.65	7/11	1/15	0/15	0/15	0/15	-
5.36	3/13	0/15	0/15	0/15	0/15	-
7.87	1/15	0/15	1/14	0/15	0/15	-
11.57	2/13	-	-	-	-	-